Joseph G. Petrosinelli (*pro hac vice*)
John E. Joiner (*pro hac vice*)
Williams & Connolly LLP
725 12th Street, NW
Washington, DC 20005
202-434-5000
202-434-5029 (facsimile)
jpetrosinelli@wc.com
jjoiner@wc.com

Loren H. Brown (*pro hac vice*)
DLA Piper LLP (US)
1251 Avenue of the Americas
New York, NY 10020-1104
212-335-4000
212-884-8546 (facsimile)
loren.brown@dlapiper.com

Matthew A. Holian (Cal. Bar No. 211728)
DLA Piper LLP (US)
33 Arch Street, 26th Floor
Boston, Massachusetts 02110-1447
617-406-6009
617-406-6109 (facsimile)
matt.holian@dlapiper.com

Attorneys for Defendant Pfizer Inc.

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA

IN RE: VIAGRA (SILDENAFIL CITRATE) PRODUCTS LIABILITY LITIGATION	Case No. 3:16-md-02691-RS MDL No. 2691
This Document Relates to: ALL ACTIONS	**PFIZER'S OPPOSITION TO PLAINTIFFS' MOTION TO COMPEL TECHNOLOGY-ASSISTED REVIEW ("TAR")**

The parties have been able to reach agreement on every aspect of a stipulated order regarding the discovery of electronically stored information ("ESI") except one—the methodology that Pfizer will use to search for responsive documents within its custodial files and certain other document sources. Pfizer has proposed an iterative process by which it selects, tests, and modifies search terms to identify those terms that yield high rates of responsive documents, while minimizing the wasteful review of documents that do not contain terms that are likely to yield responsive documents. Pfizer has used this methodology for more than a decade to review and produce documents in litigations involving ESI, and Plaintiffs do not dispute that such an approach is permitted by the Federal Rules of Civil Procedure and search terms are, in fact, a commonly used search methodology.

Rather than agree to Pfizer's methodology of iteratively applying and validating search terms to identify potentially responsive documents, Plaintiffs instead ask this Court to do what apparently no other court has ever done—force Pfizer, over its objection, to use predictive coding (also referred to as technology-assisted review ("TAR")) to identify and review potentially responsive documents. No legal authority supports Plaintiffs' request. In fact, the handful of courts that have considered this issue all have held that they do not have such authority and therefore declined to compel the use of predictive coding.

Even if this Court finds that it has the authority to force Pfizer to use predictive coding over its objection, it should not do so here. As an initial matter, the Court need not decide which methodology is superior. It only has to decide if Pfizer's methodology satisfies its discovery obligations. For the reasons explained below, iteratively applying and validating search terms (called "iterative search and validation") is an appropriate and reasonable methodology that satisfies Pfizer's discovery obligations under Federal Rule of Civil Procedure 26, which Plaintiffs do not dispute. Even if the Court determines that it should be the first court to hold that it has authority to force Pfizer to use predictive coding, there is no basis to conclude that predictive coding is inherently superior to iterative search and validation. Therefore, the Court should deny Plaintiffs' motion to compel.

I. NO LEGAL AUTHORITY SUPPORTS PLAINTIFFS' REQUEST.

No case law or rule supports Plaintiffs' request that the Court compel Pfizer to use predictive coding. In their brief, Plaintiffs do not cite a single case in which a court has ordered a producing party to use predictive coding over its objection, nor did Plaintiffs identify any such authority—despite multiple requests from Pfizer—in the parties' extensive negotiations about the issue. Nor could Plaintiffs, since it appears that no such authority exists. Instead, Plaintiffs cite only to cases in which the parties *agreed* to use predictive coding. *See* Pls.' Mot. to Compel 2-4; *In re Actos (Pioglitazone) Prods. Liab. Litig.*, MDL No. 6:11-md-2299, 2012 WL 7861249, at *1 (W.D. La. July 27, 2012) (entering ESI order "pursuant to the agreement reached between the Plaintiffs and Defendants"); *Moore v. Publicis Groupe*, 287 F.R.D. 182, 183 n.1 (S.D.N.Y. 2012) (noting that "the Court did not order the parties to use predictive coding" and "[t]he parties had agreed to defendants' use of it"); *In re: Bair Hugger Forced Air Warming Prods. Liab. Litig.*, MDL No. 15-2666, 2016 WL 3702959, at *1 (D. Minn. July 8, 2016) (entering ESI order "pursuant to the Parties' agreement"). None of these cases addresses the issue here: whether a court can compel a producing party to use predictive coding over its objection.

Also absent from Plaintiffs' motion is any meaningful discussion of the handful of cases that have addressed this issue. That is likely because every court that has done so has refused to compel a producing party to use predictive coding. *See Hyles v. New York City*, No. 10-cv-3119, 2016 WL 4077114, at *1 (S.D.N.Y. Aug. 1, 2016) (denying plaintiff's motion to compel defendant to use TAR); *In re Bridgeport Educ., Inc. Securities Litig.*, No. 12-cv-1737, 2014 WL 3867495, at *4 (S.D. Cal. Aug. 6, 2014) (denying plaintiffs' request that defendants use predictive coding on documents already produced); *In re Biomet M2a Magnum Hip Implant Prods. Liab. Litig.*, No. 3:12-md-2391, 2013 WL 1729682, at *3 (N.D. Ind. Apr. 18, 2013) (denying plaintiffs' motion to compel defendants to use predictive coding on documents already produced); *Kleen Prods. LLC v. Packaging Corp. of Am.*, No. 10-cv-5711, 2012 WL 4498465, *5 (N.D. Il. Sept. 28, 2012) (denying plaintiffs' motion to compel defendants to use content-based advanced analytics and ordering the parties to meet and confer regarding modifications to

defendants' search terms); *see also Rio Tinto PLC v. Vales S.A.*, 306 F.R.D 125, 127 n.1 (S.D.N.Y. 2015) ("[W]here the requesting party has sought to force the producing party to use TAR, the courts have refused.").

Hyles v. New York City, a case that Plaintiffs cite only in passing in a footnote, is directly on point. *See* Pls.' Mot. to Compel at 4 n.2. In *Hyles*, Magistrate Judge Andrew Peck, a leading authority on e-discovery and a self-proclaimed "judicial advocate for the use of TAR," considered whether "the defendant City (i.e., the responding party) can be forced to use TAR (technology assisted review, aka predictive coding) when the City prefers to use keyword searching." 2016 WL 4077114, at *1, 3. Magistrate Judge Peck held: "The short answer is a decisive 'NO'." *Id*. at *1.

Similar to this case, in *Hyles*, the issue of whether a court can compel a producing party to use predictive coding arose "before the City spent much, if any, money on searching for responsive ESI[.]" *Id*. at *2. The court relied on the Sedona Principles,[1] in particular Principle 6, to reach its conclusion that "the Court cannot, and will not, force the City to [use predictive coding]." *Id*. at *3. Sedona Principle 6 states: "Responding parties are best situated to evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information." *Id*. (quoting The Sedona Principles: Second Edition, Best Practices Recommendations & Principles for Addressing Electronic Document Production, Principle 6 (2007)). As the court explained, "the City as the responding party is best situated to decide how to search for and produce ESI[.]" *Id*. In evaluating the City's decision, Magistrate Judge Peck explained, "the standard is not perfection, or using the 'best' tool, but whether the search results are reasonable and proportional." *Id*. Magistrate Judge Peck recognized that keyword searches are an acceptable methodology and that the Court did not have the power to

[1] The Sedona Principles are a set of electronic document production guidelines established by The Sedona Conference Working Group, a forum for lawyers, consultants, academics, and jurists to address current problems in the areas of antitrust law, complex litigation, and intellectual property. *See* The Sedona Conference Working Group Series, www.thesedonaconference.org (last visited Sept. 22, 2016).

"force the responding party to use TAR." *Id*. at *2. He noted, "There may come a time when TAR is so widely used that it might be unreasonable for a party to decline to use TAR. We are not there yet." *Id.* at *3. Thus, because no legal authority supports compelling Pfizer to use predictive coding, the Court should deny Plaintiffs' motion.

II. ITERATIVE SEARCH AND VALIDATION IS APPROPRIATE AND REASONABLE.

Even if there were precedent for Plaintiffs' request that the Court order Pfizer to use predictive coding, the Court need not decide which method is superior to reach that issue. Rather, the Court need only decide whether Pfizer's methodology is sufficient under the Federal Rules of Civil Procedure. Indeed, courts consistently have declined to force a producing party to use predictive coding for a reason: the Federal Rules of Civil Procedure require only that a party perform a reasonable search for responsive documents, the scope of which must be "proportional to the needs of the case." Fed. R. Civ. P. 26(b)(1). Proportionality places "reasonable limits on discovery," the key to which "is careful and realistic assessment of actual need." Chief Justice John Roberts, 2015 Year-End Report on the Federal Judiciary at 6-7, available at https://www.supremecourt.gov/publicinfo/year-end/2015year-endreport.pdf. "The issue . . . isn't whether predictive coding is the better way of doing things," it is whether Pfizer's methodology "satisfies its discovery obligations," *i.e.*, is reasonable and proportional. *In re Biomet*, 2013 WL 1729682, at *2; *see also Hyles*, 2016 WL 4077114, at *3 ("the standard is not perfection, or using the 'best' tool, but whether the search results are reasonable and proportional"). Here, there is no evidence that Plaintiffs have an "actual need" for Pfizer to use predictive coding as opposed to more traditional search terms, or that Pfizer's methodology is unreasonable.

While Plaintiffs express a preference for predictive coding, they do not dispute that Pfizer's methodology satisfies its discovery obligations under Rule 26. Pfizer's methodology is an iterative process in which Pfizer tests search terms and validates them using rigorous sampling of potentially responsive documents. At the outset, Pfizer tests and validates a number of search terms by running those terms over a sample set of documents, reviewing the documents

for responsiveness, and verifying that the search terms yield high responsiveness rates. The parties then exchange lists of proposed search terms. Pfizer will agree to use any of the proposed search terms that appear on both parties' lists. Pfizer then uses sampling of the documents that do not contain the parties' agreed terms to determine the responsiveness rates of the remaining proposed search terms on Plaintiffs' list. The parties negotiate which additional search terms, if any, Pfizer will agree to use with the benefit of the responsiveness rate results, which provide information about the likelihood the terms will yield responsive documents. Often, Pfizer suggests modifying the search terms so that they yield higher responsiveness rates.[2]

This is not a novel process, either for Pfizer or other defendants. Pfizer has been using iterative search and validation successfully for more than a decade. Plaintiffs' claim that "Pfizer has routinely used TAR internally to identify and prioritize production of responsive ESI" is simply false. *See* Pls.' Mot. to Compel at 5. Pfizer never has used predictive coding to produce documents in any litigation.

More importantly, Plaintiffs have offered no evidence to suggest that iterative search and validation would not yield responsive documents. "[T]ypically, courts give deference to a producing party's choice of search methodology[.]" *Progressive Cas. Ins. Co. v. Delaney*, No. 2:11-cv-00678, 2014 WL 3563467, at *10 (D. Nev. July 18, 2014). There is no reason for this Court to depart from that deference here, when Plaintiffs do not even claim that Pfizer's methodology is unreasonable. Because Pfizer has successfully used iterative search and validation for more than a decade and there is no evidence before the Court that it is unreasonable, the Court should find that it satisfies Pfizer's discovery obligations under Rule 26 and deny Plaintiffs' motion.

[2] To be clear, while the selection of search terms is an iterative process, once the parties agree on a set of search terms (with the involvement of the Court, if necessary), Pfizer applies those terms consistently across the potentially responsive documents so that it can review and produce documents once from each document source.

III. PREDICTIVE CODING IS NOT NECESSARILY SUPERIOR TO ITERATIVE SEARCH AND VALIDATION.

Finally, even if the Court decides that it has the authority to order a party to use predictive coding over its objection *and* the relevant question is whether predictive coding is superior to iterative search and validation (rather than simply whether iterative search and validation is reasonable), the Court cannot find that predictive coding is superior to iterative search and validation on the record before it. While Plaintiffs claim that predictive coding is a "more sophisticated tool" than search terms and it is "the best, most effective, efficient and economical way to identify responsive documents and ESI," these claims are inaccurate for a number of reasons. Pls.' Mot. to Compel at 2, 4.

First, the utility of predictive coding is highly dependent on the quality of the human review process. When using predictive coding, human users train a computer to recognize responsive documents. *See* Pls.' Ex. A. Based on the training, the computer develops an algorithm that can be applied to a new document collection to identify responsive documents. *Id*. Typically, an attorney familiar with the case trains the computer by reviewing a "training set" of documents and coding those documents as responsive or non-responsive. *Id*. The computer analyzes the training set and adapts its algorithm to replicate the attorney's coding decisions. *Id*. The process is dynamic, with the training set growing in size (encompassing more documents) as the algorithm develops. Because the computer "learns" which documents are responsive from the training set, however, "the effectiveness of predictive coding highly depends on the reliability of the 'seed set.'" *Id*. at 2. In other words, predictive coding is "not a magic, Staples-Easy-Button, solution appropriate for all cases." *Moore*, 287 F.R.D. at 189.

Second, Plaintiffs claim that "[s]earch terms have been shown to return only 20-24% of responsive documents while TAR can return 80% and more." Pls.' Mot. to Compel at 4. Plaintiffs cite studies by Blair & Maron, Tomlinson, and Oard to support this claim. *Id*. at 4 n.1. Blair & Maron was published in 1985. *Id*. There have been considerable advances in both the technology and methodology of using search terms in the intervening thirty years that render

Blair & Maron's findings outdated and uninformative. While Tomlinson and Oard are more recent, their findings were not based on using an iterative search and validation process. *Id*.

More informative than these three studies is a 2012 study by the Electronic Discovery Institute ("EDI") and Oracle Corporation led by researchers at Stanford University. The first phase of the EDI-Oracle study involved thirteen e-discovery vendors that each used different predictive coding methodologies to identify responsive documents in the same dataset. *See* Monica Bay, *EDI-Oracle Study: Humans Are Still Essential in E-Discovery*, N.J. LAW JOURNAL (Nov. 26, 2013), attached as Ex. 1. The researchers evaluated a number of different outcomes, including cost and success in identifying responsive documents compared to human review. *Id*. The results showed considerable variability in cost and quality among the various predictive coding methodologies, leading the researchers to conclude that "software is only as good as its operators" and "human contribution is the most significant element" in the successfulness of a search methodology. *Id*. In fact, "a single senior-level attorney who spent 64.5 hours on review and analysis . . . performed best at finding both responsive documents and privilege documents." *Id*. Thus, predictive coding is only as good as the search methodology it utilizes. Accordingly, Plaintiffs' claim that "TAR can return 80% and more" responsive documents and is superior to search terms is baseless.

Third, Plaintiffs cite the *Actos* litigation as a model for the feasibility of predictive coding, suggesting that the use of predictive coding was seamless in that litigation. Pls.' Mot. to Compel at 2-3. Not so. In *Actos*, defendants produced only 4,000 responsive documents using predictive coding. *See* Defs.' Opp. to Pls.' Mot. to Compel Production of Docs., *In re Actos (Pioglitazone) Product Liab. Litig.*, 6:11-md-02299 (W.D. La. June 19, 2013), at 1, attached as Ex. 2. However, the production of these documents was "not just, speedy, or inexpensive[.]" *Id*. As defendants explained in their opposition to plaintiffs' motion to compel the production of additional document using predictive coding, "[t]he predictive coding process in this MDL has been lengthier and more costly than Defendants anticipated." *Id*. at 3. "[T]he algorithm [was] not an efficient means of unearthing responsive documents that [had] not already been

produced," and there was no indication that the documents that were produced were legally significant. *Id*. at 4, 8. Ultimately, the parties resolved this discovery dispute without judicial intervention, but the fact that it arose demonstrates that predictive coding is not necessarily "the best, most effective, efficient and economical way to identify responsive documents and ESI." Pls.' Mot. to Compel. at 4.

Fourth, Plaintiffs claim that in order for predictive coding to be fair and effective, they need to be involved in the computer training process. Pls.' Mot. to Compel at 3. In other words, Plaintiffs' counsel want to review the training set and have the ability to second-guess the responsiveness determinations of Pfizer's counsel. For those who still recall paper document review, this is the equivalent of having Plaintiffs' counsel present at Pfizer's document storage facility and allowing Plaintiffs' counsel to watch over the shoulders of Pfizer's counsel as they review documents for responsiveness. Not only is this approach completely unnecessary (defense counsel always make responsiveness determinations without input from plaintiffs' counsel), it would allow Plaintiffs' counsel to view in the training set: (1) documents protected by the attorney-client privilege and/or work product doctrine; and (2) non-responsive documents, which may contain proprietary or commercially sensitive information, or other irrelevant information that Plaintiffs are not entitled to view. While Plaintiffs may not consider this to be a significant harm, Pfizer does. Even if the Court ordered Pfizer to use predictive coding, there is simply no reason why Plaintiffs' counsel would need to be involved in the computer training process.

As demonstrated above, many of Plaintiffs' claims about predictive coding are inaccurate. Because predictive coding is not necessarily the superior search methodology that Plaintiffs claim it to be, the Court should find that its use is not necessary in this case.

IV. CONCLUSION

The Court should deny Plaintiffs' motion to compel because there is no legal authority that supports it. Even if the Court determines that it has the authority to force Pfizer to use predictive coding over its objection, it should not do so because Pfizer's methodology is

appropriate and reasonable and satisfies Pfizer's discovery obligations under Rule 26. Contrary to Plaintiffs' assertion, predictive coding is not necessarily superior to iterative search and validation, and there is no basis for the Court to conclude that predictive coding is superior to iterative search and validation on the record before it.

Dated: September 30, 2016

Respectfully submitted,

BY: /s/ Matthew A. Holian_____

Williams & Connolly LLP
Joseph G. Petrosinelli (*pro hac vice*)
John E. Joiner (*pro hac vice*)
725 12th Street, NW
Washington, DC 20005
Telephone: (202) 434-5000
Facsimile: (202) 434-5029 (fax)
Email: jpetrosinelli@wc.com
Email: jjoiner@wc.com

DLA Piper LLP (US)
Loren H. Brown (*pro hac vice*)
1251 Avenue of the Americas, 45th Floor
New York, NY 10020
Telephone: (212) 335-4500
Email: loren.brown@dlapiper.com

DLA Piper LLP (US)
Matthew A. Holian (Cal. Bar. No. 211728)
33 Arch Street, 26th Floor
Boston, MA 002110-1447
Telephone: (617) 406-6009
Facsimile: (617) 406-6109
Email: Matt.Holian@dlapiper.com

Attorneys for Defendant Pfizer Inc.

CERTIFICATE OF SERVICE

I hereby certify that on September 30, 2016, the foregoing PFIZER'S OPPOSITION TO PLAINTIFFS' MOTION TO COMPEL TECHNOLOGY-ASSISTED REVIEW ("TAR") was filed via the Court's CM/ECF system, which will automatically serve and send email notification of such filing to all registered attorneys of record.

Dated: September 30, 2016

By: /s/ Matthew A. Holian

DLA Piper LLP (US)
Matthew A. Holian (Cal. Bar. No. 211728)
33 Arch Street, 26th Floor
Boston, MA 002110-1447
Telephone: (617) 406-6009
Facsimile: (617) 406-6109
Email: Matt.Holian@dlapiper.com